Towards a more efficient human-exoskeleton assistance

Sara Monteiro
Center for MicroElectroMechanical Systems/ LABBELS - Associative Laboratory, University of Minho
Braga/Guimarães, Portugal
a88335@alunos.uminho.pt

Joana Figueiredo
Center for MicroElectroMechanical Systems/ LABBELS - Associative Laboratory, University of Minho
Braga/Guimarães, Portugal
joana.figueiredo@dei.uminho.pt

Cristina Santos
Center for MicroElectroMechanical Systems/ LABBELS - Associative Laboratory, University of Minho
Braga/Guimarães, Portugal
cristina@dei.uminho.pt

***Abstract*—There is evidence that the energy expended by humans can be reduced by wearing lower limb exoskeletons with user-oriented assistance strategies, such as human-in-the-loop (HITL) controllers. HITL algorithms can be implemented in exoskeletons for the automatic and online optimization of controller parameters, such as the torque profile, depending on the energy expenditure (EE) measured in real-time. This way, it is possible to minimize the EE and tailor the exoskeleton assistance for each specific user. But measuring EE is not trivial. It is more commonly estimated by indirect calorimetry, however, this method requires expensive equipment, takes too long, and is infeasible for everyday use in the real world. Therefore, this study explores machine and deep learning regression models (RMs) as EE estimators in different motor activities based on data acquired by wearable sensors and anthropometric features. Several inputs were tested but the best performance was achieved by the heart rate, the 3-axis acceleration of the chest, wrist, thigh, and ankle, and the body mass index. Results from a public dataset are presented, after the preprocessing of the data. The best-performing RM was an exponential Gaussian process regressor (GPR), that obtained root-mean-squared errors of 0.56 W/kg, 0.45 W/kg, and 0.60 W/kg for the standing, sitting, and walking activities, respectively. The GPR model outperformed a support vector machine, a boosted decision tree, a bagged decision tree, and a convolutional neural network.**

***Index Terms*—artificial intelligence, energy expenditure, human-in-the-loop control, exoskeleton assistance, wearable sensors**

I. INTRODUCTION

An ongoing cause of concern in the industry is work-related musculoskeletal disorders (WMSDs), one of the most common work-related health problem in developed countries, affecting millions of European workers across all sectors and occupations [1]. Some industry activities that are critical for WMSDs in the lower limbs are standing for long periods and carrying/lifting heavy loads [1]. Roughly 63% of workers that frequently handle heavy loads reported suffering from WMSDs [1].

Lower limb exoskeletons (LLEs) can be employed for power augmentation of workers by reducing the physical stress and strain on the user's muscles while they perform motions like walking, squatting, stand-to-sit, and sit-to-stand, or while they are in stationary positions (either sitting or standing) [2]. However, current exoskeletons can not yet provide ideal and individualized assistance to each user, raising the need for smarter closed-loop control schemes able to adapt the exoskeleton assistance to the human's needs [3, 4, 5].

One possible strategy is the real-time optimization of the exoskeleton's control parameters based on physiological signals obtained from the user - human-in-the-loop (HITL) control. This strategy shows potential for industrial applications, as it could be used to minimize the fatigue of workers while they perform heavy tasks [3]. HITL controllers currently being developed focus on the continuous adaptation of the exoskeleton torque profiles, in one or more joints, to minimize the energy expenditure (EE) of its users [3, 4, 6]. However, the use of HITL controllers has been limited to clinical applications, and no study has exploited its potential for industrial assistance yet.

The standard method to estimate the EE is indirect calorimetry, requiring a respirometer device to measure the consumption of oxygen and production of carbon dioxide [7]. However, this method uses expensive equipment, is very time-consuming, and is not practical for real-world applications [4]. To overcome this drawback, machine [8, 9] or deep [10, 11] learning regression models (RMs) have been developed to estimate the EE based on data obtained from wearable sensors. Despite the progress in RMs for estimating EE, it remains unclear which is the best architecture (input estimators and regressor), considering a trade-off between the minimal estimation error and the minimal number of sensors for practical use in the real world. To ensure the practicality of these methods and to minimize the interference to workers' movements, it is crucial to reduce the number and size of the sensors used to estimate the EE in real-world applications.

This work proposes a HITL control strategy towards more

This work was supported in part by the Fundação para a Ciência e Tecnologia (FCT) under the Stimulus of Scientific Employment with the grant 2020.03393.CEECIND, under the national support to R&D units grant, through the reference project UIDB/04436/2020 and UIDP/04436/2020, and by the FEDER Funds through the COMPETE 2020—Programa Operacional Competitividade e Internacionalização (POCI) and P2020 with the Reference Project SmartOs Grant POCI-01-0247-FEDER-039868.

efficient human-LLE assistance during industry activities. Furthermore, it aims to present an accurate and practical RM to estimate EE in industry activities (standing, sitting, and walking) by comparing machine and deep-learning RMs and studying the best input estimators among wearable sensors and anthropometric data. In the future, the best RM will be implemented in a HITL control strategy of an LLE.

II. Materials and Methods

A. Human-in-the-loop Controller

This sub-section presents the HITL control strategy that will be developed to minimize the user's EE user during industry activities. It follows the principles of HITL control implemented for other purposes [3, 4, 5, 6]. Figure 1 shows how each stage of this work is integrated into the controller, namely the training and comparison of different RMs, the EE estimation by the best RM, the HITL optimization, the exoskeleton assistance, and the data acquisition.

Firstly, five RMs were trained and validated using a public dataset [12], by using anthropometric and wearable sensor data as the input estimators and the indirect respirometry data as the ground truth of EE. The models tested were a convolutional neural network (CNN) with a regression layer, a Gaussian process regressor (GPR), a boosted decision tree (BoDT), a bagged decision tree (BaDT), and a support vector machine (SVM). The best-performing model is then applied to estimate the user's EE in real-time based on data acquired from wearable sensors, namely the heart rate (HR) and 3-axis (3D) acceleration of the chest, wrist, thigh, and ankle.

Afterward, the estimated EE is used to optimize the exoskeleton's control parameters, namely the peak extension and flexion torque of the knee joint, with an evolutionary algorithm named covariance matrix adaptation evolution strategy (CMA-ES). The CMA-ES is the algorithm more used across the literature since it can be used for high-dimensional optimization problems [3]. This optimization algorithm finds, in real-time, the parameters of the torque reference trajectory that minimize the energy being expended by the users (cost function), while they are performing their desired industrial activities. A shape function is then used to convert the control parameters to the reference knee torque trajectory. Based on this reference and on the real torque (measured by embedded sensors in the LLE), the proportional–integral–derivative (PID) controller computes the control command that adapts the LLE's assistance to a more efficient user's condition during the industrial activity.

B. Dataset Description

The dataset [12] used in this work was first published and described by Cvetković et al. [13] and was used both by Gjoreski et al. [8] and Catal et al. [9] for EE estimation. It was chosen since it was the only publicly available dataset whose participants performed activities commonly performed by industry workers, namely standing, sitting, and walking. Additionally, this dataset contains HR and acceleration data (measured in four different human body locations), both signals that are highly correlated with EE [14, 15].

The dataset included sensor and anthropometric data from ten participants (8 male and 2 female), with ages between 24 and 33 years old ($27 \pm 3.1yr$), body mass between 64.2 kg and 101 kg ($78.4 \pm 10.9kg$), and body mass index (BMI) between 20 kg/m^2 and 28.9 kg/m^2 ($24.1 \pm 2.4kg/m^2$). The

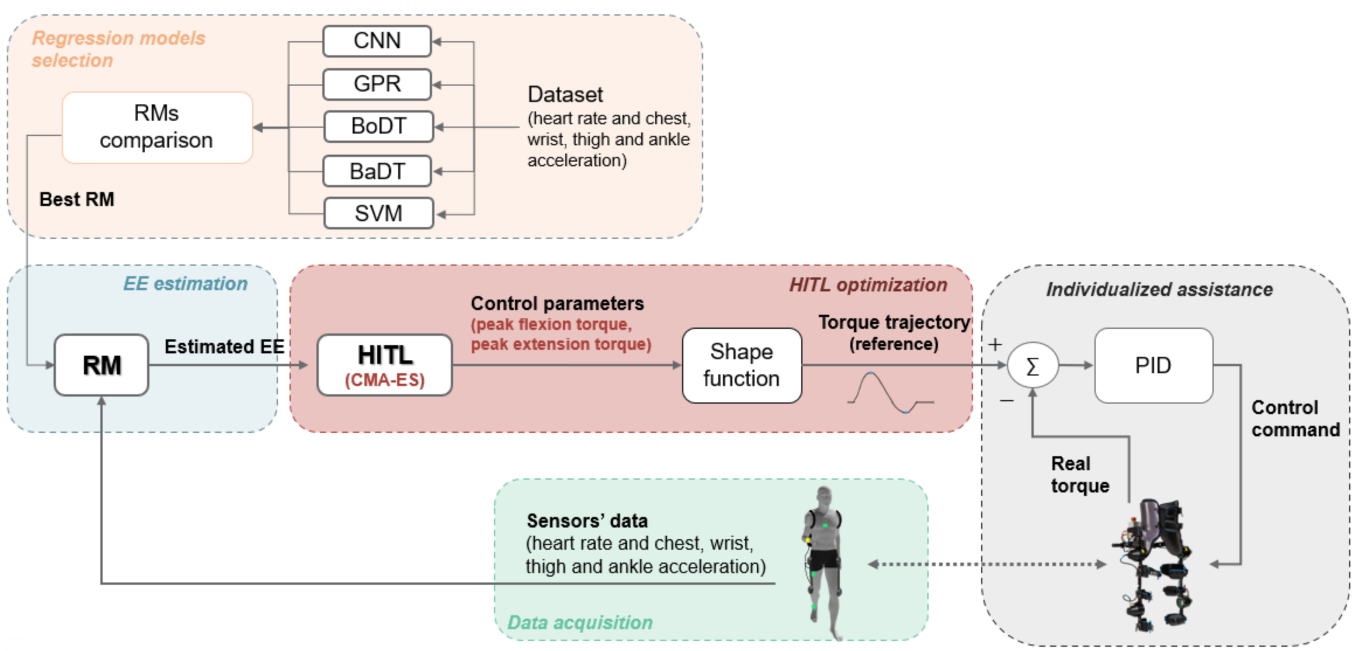


Fig. 1: Proposed HITL strategy.

participants performed 15 different activities in 13 distinct scenarios for a total of 105 minutes on average [7, 13].

They wore seven different sensors: (i) four 3D accelerometers on the chest, right wrist, right thigh, and right ankle (Shimmer 2, Ireland); (ii) one chest strap (Zephyr Bio-harness, U.S.A.) used to measure the HR, breath rate (BR), skin temperature (ST), and R-R interval (RR); (iii) one activity armband (Bodymedia Fit, U.S.A) used to measure the galvanic skin response (GSR), near body temperature (NBT), ST, and estimated EE (EEE); (iv) one smartphone on the left trouser pocket used to measure the 3D acceleration; and (v) a portable respirometer (Cosmed k4b2, Italy) used to measure the oxygen consumption (VO2) and carbon dioxide production (VCO2). Figure 2 shows the location of these sensors, the measured signals, and their sampling frequency [7].

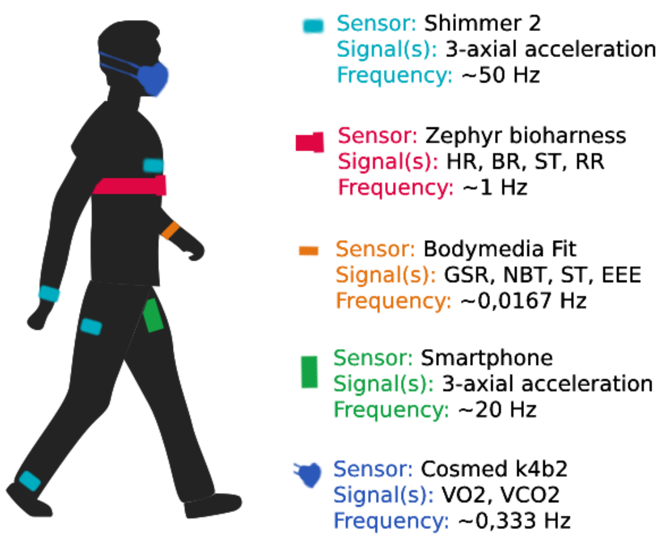


Fig. 2: Representation of on body location of the sensors, measured signals, and the corresponding acquisition frequency.

From the original dataset, we selected four variables of interest, namely the 3D acceleration, HR, BMI, and EE, along with the activities relevant to this study: (i) standing; (ii) transitioning down; (iii) sitting; (iv) transitioning up; (v) walking. The data was organized into three separated sub-datasets: standing (i), sitting (ii, iii, iv), and walking (v). Furthermore, it was noticed that the dataset did not include x-axis acceleration data for participant I (9th participant). This was not mentioned nor explained in any of the articles [7, 8, 13]. This led to the exclusion of participant I's data from the sub-datasets.

C. Data Preprocessing

The following preprocessing methods were applied to each sub-dataset. First, the EE data was converted from metabolic equivalent of task (MET) to Watt (W). Then, both the HR and EE were normalized by the body mass of each participant [11]. The effect of this normalization on the RMs performance was studied. Further, the acceleration signals were filtered with a real-time zero-phase bidirectional 4th-order Butterworth [8, 16]. Two filters were compared: (i) a low-pass filter with a cut-off frequency of 20 Hz and (ii) a band-pass filter with low cut-off frequencies of 0.1 Hz and a high cut-off frequency of 20 Hz.

Subsequently, the input estimators (acceleration and HR) and the EE were segmented into 10-second windows, and the windows were tested with and without overlaps of 5 seconds, used to enable faster updates of the EE estimation in real-time [8, 9]. For each 10-second segment, we computed the data average and subtracted it by its average during the rest period, given by the average of each signal during the first 15 minutes of the dataset, obtained while each participant was laying down. The length of the windows was chosen as 10 seconds since during this period, on average, a subject performs 2 to 3 breath cycles, which may be enough for estimating the instantaneous EE.

After data segmentation, we balanced each sub-dataset to the same amount of data from each participant. During this process, we removed the data of participant J from the walking sub-dataset given the lower number of samples when compared to the other participants. Furthermore, given the different sampling rates of each sensor, we applied an interpolation method (piecewise cubic interpolation) in the HR and EE signals [10, 17].

Moreover, the sub-datasets were analyzed to extract outliers from the sub-datasets. By studying the average and standard deviation of EE for the walking, sitting, and standing sub-datasets, it was noticed that participant G presented an average EE higher than the average of all participants by two times the standard deviation, in the standing sub-dataset. Therefore, participant G was removed from this dataset. Figure 3 presents a flowchart of the implemented preprocessing steps. Finally, the input estimators and the EE ground truth were normalized before the model's training. We compared three normalization methods: median normalization, min-max normalization, and z-score normalization.

D. Models' description

The RMs were implemented using a team-owned deep-learning regression tool and the Regression Learner App, both implemented in Matlab (2022a, The Mathworks, Natick, MA, USA). The deep-learning regression tool was used to train CNNs with different architectures and hyperparameters. Each CNN was composed of one to three convolutional layers, each followed by a rectified linear activation function (ReLU) layer and a max pooling layer with a pool size and stride of 2. After the convolutional layer(s), a global average pooling layer was implemented, followed by a fully connected (FC) layer and, at last, a regression layer. The model's hyperparameters were optimized for EE estimation, namely the number of filters and their size on each convolutional layer, the number of hidden neurons on the FC layer, the learning rate, and the batch size. Table I summarizes the hyperparameters studied for the CNN. We used the Regression Learner App to build the following

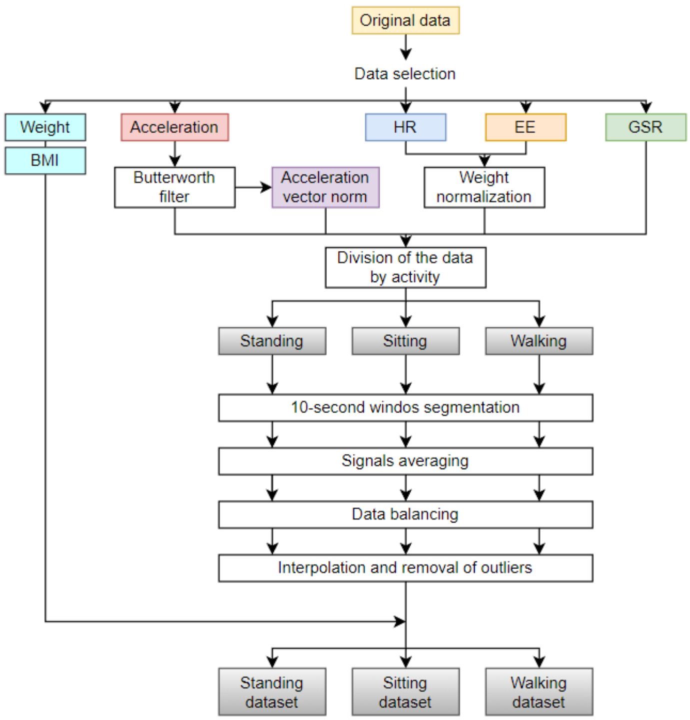


Fig. 3: Flowchart of the implemented preprocessing steps.

machine learning regressors: a GPR (with an exponential isotropic kernel function), a BoDT (with a minimum leaf size of 8 and 30 learners), a BaDT (with a minimum leaf size of 8 and 30 learners), and an SVM (with a linear kernel function, epsilon of 0.15, and a C parameter of 1.5).

TABLE I: Tested hyperparameters for the CNN

Hyperparameter	**Values**
Number of convolutional layers	1, 2, or 3
Number of filters	8 to 360
Filter size	5 to 30
Hidden neurons	100 to 1000
Learning rate	0.01, 0.005, or 0.001
Batch size	8 to 360

E. Models' Training and Evaluation

For training the RMs, we studied the effect of two sets of input estimators: (i) 3D acceleration, HR, and BMI; (ii) 3D acceleration, acceleration vector norm, HR, GSR, BMI, and body mass. This study will verify if the addition of the acceleration vector norm, GSR, and body mass increases the RM's performance.

From a visual inspection of HR and EE, we noticed that these measurements present a high variability across activities. We hypothesized that EE could be better estimated if three different RMs were created to predict the EE during each activity, separately. To study this hypothesis, two different approaches were tested. In the first approach, the data was all given to the same RM and the activity type was provided as an input to the models. In the second approach, three different RMs were trained with each sub-dataset (standing, sitting, and walking), and, therefore, the activity type was not given as input to the models.

Before the models' training, each sub-dataset was divided into training/validation data and test data. Then, the training data was shuffled. The validation method enforced was leave-one-out cross-validation (LOOCV), a form of *k*-fold cross-validation where *k* (the number of folds) is equal to the number of subjects used for training (7 for the standing and walking sub-datasets, and 8 for the sitting sub-dataset). During *k*-fold cross-validation, the training dataset is divided into *k* sets, where *k-1* are used to train the model and the other set is used to validate it. This process is repeated *k* times, meaning that each subject was used for validation once.

The performance of each RM was obtained, for each of the validation iterations of the LOOCV method, and, in the end, the average and standard deviation of each metric was computed. The models were evaluated regarding the metric obtained during the validation process, namely the root-mean-squared error (RMSE), which measures the difference between the predicted values and the target values, and the coefficient of determination (R^2), which assesses the fit quality of the RMs.

III. RESULTS

To achieve the best possible results and obtain a robust RM capable of estimating the EE with the lowest possible error, various tests were performed.

A. Preprocessing methods

Regarding the data preprocessing steps, we studied the following aspects: (i) the filtering of the acceleration signals, comparing the effects of low-pass and band-pass Butterworth filters; (ii) the effect of normalizing the HR and EE to the user's body mass; (iii) the time window of data segmentation, with and without the 5-second overlaps; and (iv) the normalization method.

The filter that led to better model performance was the low-pass filter. The model also performed better when the HR and EE were normalized by the user's body mass, compared to the use of body mass as an input estimator. The results also indicated that the use of 5-second overlaps and averaging the EE in the 10-second window, and the use of median normalization allowed for better regression results.

B. Model's Inputs

Results of the models' validation indicate that the inclusion of additional inputs (acceleration vector norm, GSR, and body mass) did not decrease the estimation error. Based on this finding and the need of minimizing the number of estimators for real-time EE estimation, the selected inputs were the 3D acceleration, HR, and BMI.

C. General vs Activity-specific Models

The results showed that the use of the activity-specific RMs led to better performance during the models' validation when compared to a general model capable to estimate the EE during all three activities.

D. Best Convolutional Neural Network

Figure 4 presents the CNN architecture that presented the best results during the LOOCV method and its best hyperparameters.

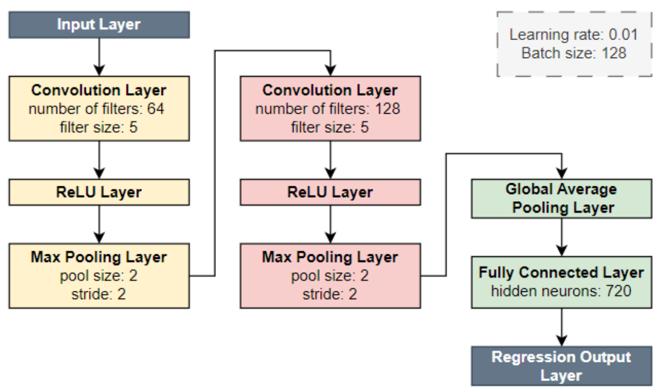


Fig. 4: Architecture and hyperparameters of the best-performing CNN.

E. Models' Comparison

Table II presents the performance of the best CNN and each machine-learning model during the validation phase, for each activity. These results showed that the model with the best performance (lowest RMSE and higher R^2), for every activity type, was the GPR.

TABLE II: Validation results of the studied models, namely the mean and standard deviation values of the RMSE and R^2

	RMSE (W/kg)			R^2		
	Standing	Sitting	Walking	Standing	Sitting	Walking
CNN	0.93 ± 0.11	0.64 ± 0.11	1.27 ± 0.38	0.36	0.67	0.69
GPR	0.55 ± 0.07	0.45 ± 0.07	0.60 ± 0.04	0.83	0.70	0.84
BoDT	0.71 ± 0.07	0.54 ± 0.05	0.74 ± 0.08	0.71	0.58	0.76
BaDT	0.65 ± 0.05	0.51 ± 0.06	0.68 ± 0.06	0.75	0.61	0.80
SVM	0.92 ± 0.09	0.67 ± 0.05	0.82 ± 0.04	0.51	0.34	0.70

Figure 5 shows the predicted EE during validation and its comparison to the real EE, for each studied activity. The figure also depicts the ideal model's performance, represented by an identity line (prediction equal to the target).

IV. Discussion

This study contributes by presenting a HITL control strategy for minimizing workers' EE while performing industrial activities. Furthermore, it studies the effect of the input estimators, preprocessing methods, and RMs towards achieving a more accurate RM for estimating EE.

The results showed that the preprocessing methods used to prepare the data can affect the efficacy of an RM. Better results were achieved by using a low-pass filter to remove high-frequency noise from the acceleration since the band-pass filter ended up erasing small variations in these signals that were related to changes in activities. A low-pass filter was also used by some studies in the literature [9, 11, 17]. The 5-second overlaps were useful to increase the data inserted in the RMs, as verified by the results. Normalizing the data by the participants' body mass, which was also performed by some studies [11, 14], improved the model's learning capability as well, demonstrating the impact of a person's body mass on his/her EE.

Furthermore, it was concluded that providing more inputs to the models did not improve their performance, since the estimator signals (acceleration vector norm, GSR, and body mass) did not correlate as much to changes in the participants' EE. Previous studies have also indicated that acceleration and HR are highly correlated to EE [14, 15]. It was also observed that activity-specific RMs were more capable of estimating the EE than a general model.

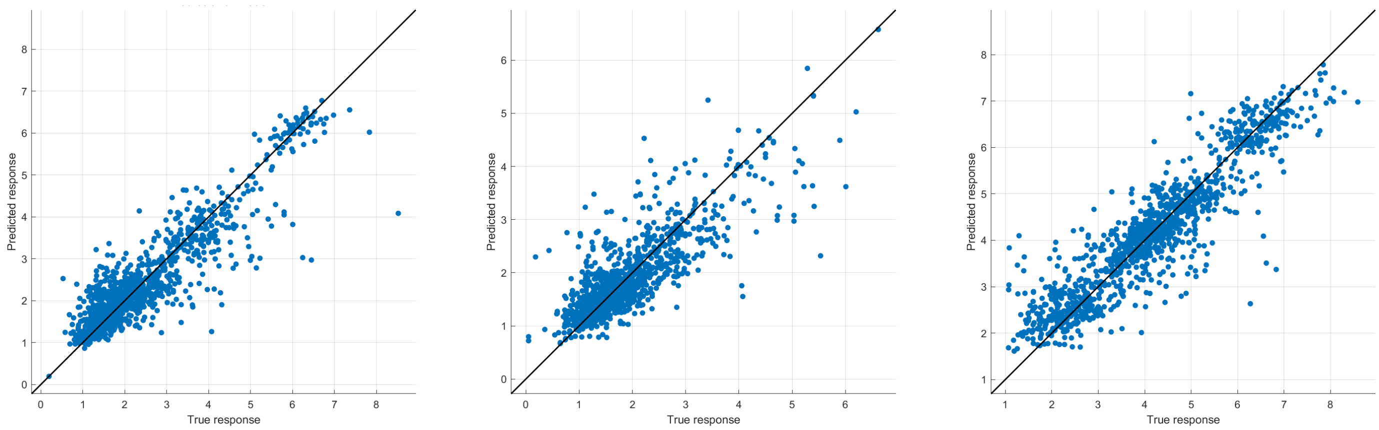
During the LOOCV, the RMs studied obtained errors in the order of magnitude of the errors achieved by similar studies in the literature (RMSEs of 0.36 W/kg for the walking activity [8] and 1.03 W/kg for walking, running, cycling, and ascending stairs [14]). For the model comparison, we verified that the GPR outperformed the remaining machine learning models and the CNN deep learning model. In fact, most of the literature studies have shown that machine learning models can successfully estimate EE, such as SVMs [8][16] and BoDTs [9]. In regards to the studies performed on the same dataset, Gjoreski et al. [8] obtained better performance with a regression SVM, while Catal et al. [9] achieved better results with a BoDT.

The performance achieved by the GPR model was similar to the performances obtained by the studies that used the same dataset [8, 9], despite using significantly fewer sensor data. Therefore, this work innovates by creating a EE estimating model feasible for HITL applications without considerable changes in efficacy, resulting in less obtrusive exoskeleton assistance for workers during occupational tasks.

V. Conclusions

This study presents first the architecture of a HITL control strategy to minimize the user's EE (i.e., to maximize his/her metabolic efficiency) through LLE's assistance. It also studied the relevant preprocessing methods, input estimators, and RMs to estimate EE. The results showed that a simple machine-learning model, with access to minimal data from only five wearable and light sensors, can produce promising results, close to what is observed in the literature. This demonstrated the reliability of using RMs to estimate EE in real-time and in the real world for HITL assistive strategies with LLEs.

In the future, the model will be optimized by training it with more balanced data. The results presented here will be used to integrate the selected RM into a HITL controller on an LLE. This control strategy is aimed to minimize the energy expenditure of industry workers.



(a) Standing (b) Sitting (c) Walking

Fig. 5: Predicted EE compared to the real EE, obtained during LOOCV, for the different activities. The straight line represents the ideal performance.

References

[1] J. de Kok, P. Vroonhof, J. Snijders, G. Roullis, M. Clarke, K. Peereboom, P. van Dorst, and I. Isusi, "Work-related musculoskeletal disorders: prevalence, costs and demographics in the eu," 2019.

[2] B. Chen, H. Ma, L. Qin, F. Gao, K. Chan, S. Law, L. Qin, and W. Liao, "Recent developments and challenges of lower extremity exoskeletons," *Journal of Orthopaedic Translation*, vol. 5, pp. 26–37, 2016.

[3] J. Zhang, P. Fiers, K. Witte, R. Jackson, K. Poggensee, C. Atkenson, and S. Collins, "Human in the loop optimization of exoskeleton assistance during walking," *Science*, vol. 356:6344, pp. 1280–1283, 2017.

[4] J. Koller, D. Gates, D. Ferris, C. Remy, and A. Arbor, "'body-in-the-loop' optimization of assistive robotic devices: A validation study," *Robotics: Science and Systems*, vol. 12, 2016.

[5] S. Song and S. Collins, "Optimizing exoskeleton assistance for faster self-selected walking," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 786–795, 2021.

[6] G. Bryan, P. Franks, S. Song, R. Reyes, M. O'Donovan, K. Gregorczyk, S. Collins, "Optimized hip-knee-ankle exoskeleton assistance reduces the metabolic cost of walking with worn loads," *Journal of NeuroEngineering and Rehabilitation*, vol. 18:1, 2021.

[7] J. Alvarez-Garcia, B. Cvetković, and M. Lustrek, "A survey on energy expenditure estimation using wearable devices," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–35, 2020.

[8] H. Gjoreski, B. Kaluža, G. Matjaž, R. Milić, and M. Luštrek, "Context-based ensemble method for human energy expenditure estimation," *Applied Soft Computing*, vol. 37, pp. 960–970, 2015.

[9] C. Catal and A. Akbulut, "Automatic energy expenditure measurement for health science," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 31–37, 2018.

[10] J. Zhu, A. Pande, P. Mohapatra, and J. Han, "Using deep learning for energy expenditure estimation with wearable sensors," *17th International Conference on E-health Networking, Application Services (HealthCom)*, no. 7454554, pp. 501–506, 2015.

[11] J. Lopes, J. Figueiredo, P. Fonseca, J. Cerqueira, J. Vilas-Boas, and C. Santos, "Deep learning-based energy expenditure estimation in assisted and non-assisted gait using inertial, emg, and heart rate wearable sensors," *Sensors*, vol. 22, 2022.

[12] "JSI dataset." https://dis.ijs.si/ami-repository/datasets/chiron.rar. [accessed on December 2022].

[13] B. Cvetković, R. Milić, and M. Luštrek, "Estimating energy expenditure with multiple models using different wearable sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 1081–1087, 2016.

[14] K. Ingraham, D. Ferris, and C. Remy, "Evaluating physiological signal salience for estimating metabolic energy cost from wearable sensors," *Journal of Applied Physiology*, vol. 126, pp. 717–729, 2019.

[15] A. Lucena, J. Guedes, M. Vaz, L. Silva, D. Bustos, and E. Souza, "Modeling energy expenditure estimation in occupational context by actigraphy: A multi regression mixed-effects model," *International Journal of Environmental Research and Public Health*, vol. 18, 2021.

[16] S. Su, B. Celler, E. Ambikairajah, and A. Savkin, "Estimation of walking energy expenditure by using support vector regression," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, vol. 7, no. 1617240, pp. 3526–3529, 2005.

[17] Z. Ni, T. Wu, T. Wang, F. Su, and Y. Li, "Deep multi-branch two-stage regression network for accurate energy expenditure estimation with ecg and imu data," *IEEE Transactions on Biomedical Engineering*, vol. 69, pp. 3224–3233, 2022.